

Context aware, multimodal, and semantic rendering engine

Patrick Salamin, Daniel Thalmann, and Frédéric Vexo
VRLab*- EPFL, Switzerland

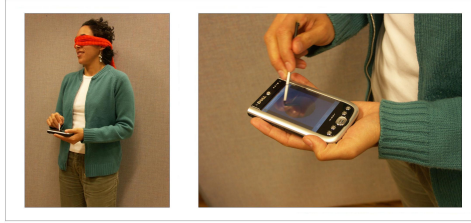


Figure 1: An example of trans-modal rendering application selected by our engine to render 3D digital content to visually impaired people.

Abstract

Nowadays, several techniques exist to render digital content such as graphics, audio, haptic, etc. Unfortunately, they require different faculties that cannot always be applied, e.g. providing a picture to a blind person would be useless.

In this paper, we present a new multimodal rendering engine with a server web-connected to other devices to perform ubiquitous computing. In order to take advantage of user capabilities, we defined an ontology populated with the following elements: user, device, and information. Our system, with the help of this ontology, aims to select and launch automatically a rendering application.

Several test case applications were implemented to render shape, text, and video information via audio, haptic, and sight channels. Validations demonstrate that our system is flexible, easily extensible, and shows promise.

CR Categories: H.5.2 [Information interfaces and presentation]: User interfaces—evaluation/methodology; Input devices and strategies; Interaction styles; User interface management systems; User-centered design; H.1.2 [Models and principles]: User/Machine systems—Human information processing; I.3.6 [Computer graphics]: Methodology and techniques—ergonomics; interaction techniques; K.4.2 [Computers and society]: Social issues—Handicapped persons/special needs;

Keywords: trans-modality, semantic rendering, context awareness

1 Introduction

Rendering digital content is not a trivial process. Several parameters need to be taken into account: user capabilities (available senses and limbs), type of content (3D, text, audio, etc), and available devices to transmit the information. Digital content rendering is then quite a hard process that often necessitates a manual adaptation of the content. Hopefully, semantics can help us to make the

content compliant to the user context (and also to allow ubiquitous computing). And the ontology gives the possibility to a system to reason on the inputs information.

This is why we implemented a multimodal, context aware, and semantic rendering engine based on ontology. Our system takes into account the three components we judge as mandatory to render efficiently 3D content: information type, available devices, and user context. User context and content advise our system about the available limbs, senses, special requirements such as mobility...and available devices. Our engine can then firstly evaluate which devices would fit to transmit the content depending on the user context. Secondly, it can adapt the content by using semantic to select a rendering solution.

In this paper, we first make a brief overview of related works within the domains of context awareness, trans-modality and semantic rendering. It will be followed by descriptions of the concept, our engine, and the test case applications. We will then present the validation experimentations to conclude with a discussion on their results and the possible further works.

2 Related works

Human-computer interaction [Jaimes and Sebe 2007] (HCI), surveyed in [Dix et al. 1997], is a not-so-old trend claiming that science and computers could better serve human needs [mar 1991]. It is then no more the human who must adapt to technology. HCI leads to the ambient intelligence [Aarts 2004] that consists in an electronic environment that is sensitive and responsive to the presence of people and their behavior. We can also notice that human-computer interaction trend contributed to the user interface improvement. For example, Kristina et al. proposed a new way to design and evaluate intelligent user interfaces in [Höök 1998]. Indeed, with the techniques improvements, the interface has to adapt to user (instead of the contrary). In this sense, user interfaces became multimodal [Corradini et al. 2003] to make inputs easier. Malkawi et al. also claimed the interest of using multimodal human-computer interaction for immersive visualization such as speech-gesture recognition and augmented reality for indoor environments[Malkawi and Srinivasan 2004].

A framework to rapidly develop multimodal interfaces was even developed in 2003 [Flippo et al. 2003]. But some modalities seem to be preferred because of their intuitiveness. For example, Cohen et al. mainly focused on tangible interfaces [Cohen and McGee 2004] and voice [Cohen and Oviatt 1994] to improve communication between humans and machines. Other researchers preferred to

*url: <http://vrlab.epfl.ch>

work with face and body gesture recognition [Gunes et al. 2004] for a vision-based multimodal analyzer while Toshiyuki developed real-time gesture recognition by learning and selective control of visual interest points application in 2005 [tos 2005]. We can also cite Pernilla et al. [Qvarfordt and Zhai 2005] and Linda et al. [Sibert and Jacob 2000] who focused on eye-gaze to allow the computer conversing with the user.

Nevertheless, some people cannot get information from every channel. There is then a need to trans-code the information for impaired users. Some very promising researches were done within the trans-coding domain, especially with the auditory and visual modalities. Jacquemin et al. e.g. developed a system involving audio- and visuo-rendering to mimic gusts of wind blowing a veil [Jacquemin and de Laubier 2006]. Within the domain of trans-modality, we also developed in 2007 an application allowing blind people to recognize 3D digital content with the help of a Smartphone [Salamin et al. 2007] to prove the efficiency of such a combination of channels (audio and touch compensate the sight).

Notice that content trans-coding process can be simplified by using semantics [Rautek and Bruckner 2007]. Knowing user context [Yoon et al. 2007][Mowafi and Zhang 2007][Kaltz 2006] is then required in order to choose the transmission channel with a user-centered interface approach. With the apparition of the “user context awareness” concept, the system can react according to the user context instead of requiring an interaction with the user such as pushing a button. User context improves then the quality and intuitiveness of the interface [Couderc and Kermarrec 1999][Martinez-Ruiz et al. 2008] but also the level of service. By the way, trans-coding the content allows making the interface personalized and thus more adapted to the user [Gui et al. 2009]. Creation of user profiles to recognize users with the help of ontology [Sutterer et al. 2008] results from this trend.

Shamsfard et al. also claimed that the most relevant tool seems to be the ontology [Shamsfard and Abdollahzadeh Barforoush 2003] to allow a system reasoning with several inputs. Indeed, using databases only allows direct and obvious links between elements instead of an ontology in which each elements has properties, requirements and more subtle links such as the evolution of some parameters depending on other variables. For example, driving a car requires sight, which means that user video channel must be considered as disabled during this action even if the user can see. Moreover, ontology seems to perfectly suit to an open and extensible system that must integrate the information from a diverse range of sources. Indeed, ontology potentially provides a well-founded mechanism for the representation of such structured information [Ye et al. 2007]. Finally, there already exists an integrated methodology of ontology engineering from scratch, inspired by various scientific disciplines, in particular database semantics and natural language processing [Spyns et al. 2008].

Based on the previous researches, we decided to implement a context aware, multimodal, semantic, and ontology-driven rendering engine. Context awareness helps the system to evaluate the user requirements, while semantic gives information about how to render it. Finally, ontology helps to find an adequate way to render the content to the user, with the help of trans-modality in some cases.

3 Concept

As written in the previous section, it is no more to the user to adapt to the application. With the current technologies, the system can be aware of the user context and provide the information via a modality adapted to his/her needs. Moreover, semantic contributes to a more efficient rendering. We need then to use an ontology describing

interactions between human and content in a given context with descriptors such as context, user capabilities, and available devices.

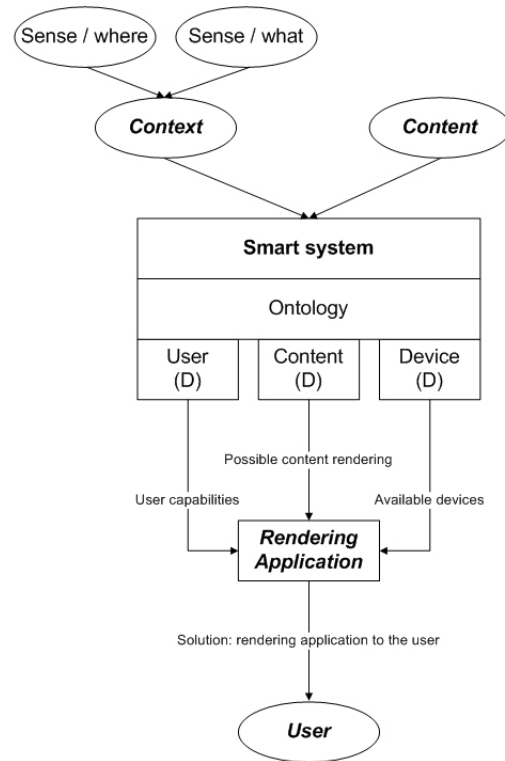


Figure 2: *Presentation of the system.*

This is why we created a multimodal rendering engine (see Fig. 2), based on such an ontology, which is aware of the type of content to transmit, the available devices, and the user content. Several test-case applications (see Fig. 3) were also implemented to render different type of information in multimodal ways. We will present them more in detail in the following sections, but we can already cite some of them: several systems providing an ex-centric perspective to the user [Salamin et al.][Salamin et al. 2008a][Salamin et al. 2009a], an application helping blind people to “visualize” 3D digital content [Salamin et al. 2007], a way to tele-rehabilitate shoulders and elbows for physiotherapists [Salamin et al. 2008b], to tele-operate a blimp with video feedback for video-surveillance [Salamin et al.], and to make that trans-coded RSS feeds follow you even if you have to leave your office [Salamin et al. 2009b]. Our engine combines then the context information with semantic rendering to automatically select one of the test-case applications. Finally, our engine can launch the chosen application on the web-connected device.

We present in the next section how our system integrates user context, how it treats this information, and the way used to select and launch the rendering applications.

4 System presentation

Our system must be aware of several elements to select an adequate rendering application: user context, device availability, type of information. We based our rendering engine on an ontology contain-

ing these elements. Concerning the information ontology, we did not restrict the kinds of information that could be a shape, a text, a video, a sound, a taste, etc. The parameters for human and device elements of the ontology especially focused on the senses and limbs of the user. If a sense (or limb) is available for the human in a specific context, our system deduces it can be used to transmit the content with a device using this sense (or limb).

4.1 Multimodal rendering engine

As you can see on Fig. 2, our system is based on an ontology describing interactions between human and content in a specific context with the help of content, user, and devices descriptors. In Fig. 2, elements in *italic* (content, user context and rendering application) change for each case because they are the input/output of our system. Indeed, the context influences the following descriptors (in the bottom of the smart system in Fig. 2: user and devices; and the content input obviously affects the content descriptor. Our system is then composed of three main parts: input (context and content entered by hand via an interface on the main terminal), rendering engine based on an ontology (smart system, ontology, and descriptors - user, content, and device), and output (rendering application connected via Wi-Fi or BlueTooth to the main terminal) that corresponds to the application selected by our system to render the content to the user.

In our ontology, the subject is considered as a set of limbs and senses that can be available and the devices are defined by the limbs and channels they require (sight, hearing, . . . , arms, and legs). The context consists in actions of the user (e.g. driving, walking, sitting) that will change subject availabilities state; And thus this will also influence the rendering device choice.

In order to show how our engine works, we will now present a concrete example of a car driver (context) who wants to receive weather forecasts (content). In this situation, context gives pieces of information to the descriptors of our system with the help of the ontology. On one hand, the engine knows that sight channel is not available because of driving context. On the other hand, the auto-radio is an available device and the content - weather forecasts - could be rendered as a visual map, an audio- or video news-flash, etc. In this case, our system will render the content as an audio news-flash (rendering application) via the car auto-radio (device). The user sight is then not required and the renderer device is obviously available.

We will now present some rendering applications, how our engine supports them and we also highlight the advantages of our system.

4.2 Rendering applications

In this section, we present the rendering applications (see Fig. 3) used to validate our system. We will describe some test cases in which the system (show in Fig. 2) reasoning leads to select one of our developed applications.

4.2.1 Video-surveillance

A situation could be a surveillance walking team helped with a video flow provided by a camera carried on a soft blimp [Thalmann et al. 2006].

In this case, content is a video flow and context requires an embed device usable while walking or running, i.e. legs or arms are considered as disabled. The best way to propose this content to the safety guards is to use the video channel via Wi-Fi, e.g. with a Smartphone. You can find hereinafter a more precise description of this rendering application that is supported by our system.



Figure 3: Quick overview of the rendering applications: third-person perspective (3PP) on the left top; 3D content visualization for blind people on the top center; improved 3PP on the right top; surveillance system and blimp tele-operation with video feedback on the second line; intelligent switch between multiple fixed cameras to provide 3PP on the third line; tele-rehabilitation on the bottom left; and RSS feeds trans-coder on the bottom right.

The system exploits advanced Mixed and Virtual Reality technologies to create a surveillance and security system that could be also

extended to define emergency prevention plans in crowd environments. Surveillance cameras are carried by a mini Blimp which is tele-operated using an innovative Virtual Reality interface with haptic feedback. An interactive control room (CAVE) receives multiple video streams from airborne and fixed cameras. Eye tracking technology allows for turning the user's gaze into the main interaction mechanism; the user in charge can examine, zoom and select specific views by looking at them. Video streams selected at the control room can be redirected to agents equipped with a PDA. On-field agents can examine the video sent by the control center and locate the actual position of the airborne cameras in a GPS-driven map. The aerial video would be augmented with real-time 3D crowd to create more realistic risk and emergency prevention plans. The prototype shows the added value of integrating AR/VR technologies into a complex application and opens up several research directions in the areas of tele-operation, Multimodal Interfaces, simulation, risk and emergency prevention plans, etc.

4.2.2 Exo- vs. ego-centric perspective

A second one could be inspired from video games. The user needs a video flow from a camera in simulation allowing a better proprioception.

In this case, content is again a video flow and context require an embed equipment that allows to walk and to see oneself to better improve the immersion [Salamin et al.]. As immersion is very important in this case, the system will propose to use sight, e.g. with a HMD instead of a Smartphone as it is more immersive.

There exist then several possibilities concerning the way to get video signal: one mobile camera following the user or several fixed ones. Indeed, as the camera is behind the user, the exocentric perspective proposes you to see your own limbs, as in virtual reality simulations it is sometimes disturbing not to be able to see your own body. It seems to create an issue in the proprio-perception of the user who does not completely feel integrated in the environment. This perspective should be beneficial for the users. We propose to give the possibility to the people to use the first and the third-person perspective like in video games (e.g. GTA). As the gamers prefer to use the third-person perspective for moving actions and the first-person view for the thin operations, we will verify this comportment is extendable to simulations in augmented and virtual reality.

The two following situations are similar to the second one but, depending on the context, user capabilities, and requirements, it could be preferable to launch one of them. Indeed, it has been proved that Third-Person Perspective (3PP) enhances user navigation in 3D virtual environments by reducing proprio-perception issues [Salamin et al.]. Nevertheless, this approach has shown drawbacks related to occlusions and adaptation time. The improved Third-Person Perspective [Salamin et al. 2008a] (i-3PP) - does allow the user to see through his/her body in order to fix 3PP limitations, e.g. occlusions. As gamers prefer using 3PP for moving actions and the First-Person Perspective (1PP) for fine operations, we proved that this behavior is extensible to simulations in augmented and virtual reality.

Another improvement of this exocentric perspective is presented hereinafter [Salamin et al. 2009a]. Augmented reality (AR) environments are suffering from a limited workspace. In addition, registration issues are also increased by the use of a mobile camera on the user that provides a first-person perspective (1PP). Using several fixed cameras reduces the registration issues and, depending on their location, the workspace could also be enlarged. In this case of an extended workspace, it has been shown that third-person perspective (3PP) is sometimes preferred by the user. Based on the previous hypotheses, we developed a system working with several fixed cameras that can provide 3PP to a user wearing a video see-

through HMD. Our system uses an "intelligent switch" to propose our "best view" to the user, i.e. avoiding markers occlusion and taking into account user displacements.

4.2.3 "Visualization" learning for visually impaired people

Imagine now that we need to render 3D digital content (e.g. basic volumes such as a sphere, a cube, etc.) to visually impaired people.

The user context indicates that we cannot use the sight channel and the situation leads us to use low-cost devices.

There exist several ways to render the content: 3D, text, 2D pictures, . . . , touch, and audio. Our system will then look for an application requiring low-cost device proposing a combination of audio and touch: a Smartphone. We implemented an application [Salamin et al. 2007] that should allow almost everybody to "see" or at least to perceive 3D shapes. Indeed, Virtual Environments (VE) are mainly visual experiments that exclude visually impaired people. This trans-modal system works with the touch and ear channels on a Smartphone, producing different tonalities depending on the user's finger location on the screen that shows the shape.

4.2.4 Telerehabilitation

Another situation would be a physiotherapist that must perform the rehabilitation of a patient from another location.

As the user (physiotherapist) is not always at the same location, the used device must be mobile and web-connected to be linked to the rehabilitation engine (in the patient room). Moreover, the physiotherapist needs a video and audio feedback and the possibility to apply different and accurate forces onto the patient.

A haptic device is then required at the patient location, and the physiotherapist needs an intuitive solution to apply the forces, e.g. graphic and interactive (touchable) interface with video channel. We need then a mobile device that is web-connected, provides audio and video feedback and coupled with a camera. We will then use a tele-rehabilitation system aiming to help the physiotherapists for the shoulder and elbow treatment [Salamin et al. 2008b]. It is based on a two-arm haptic force feedback to avoid excessive efforts and discomfort with the spinal column. This system, remotely controlled by smart phone, has been validated by a physiotherapist with the help of muscular effort measurements (EMG).

4.2.5 Stay tuned!

A last but not least case could be a user who must go to an appointment by car while he/she was reading RSS feeds on the desktop computer screen [Salamin et al. 2009b].

In this case, the user is driving a car, which means that we cannot use sight channel. But, as we are in a car, there is an auto-radio that could be the rendering device because it proposes content via the audio channel. Concerning the content, a feed could be provided as a video, audio news flash, a visual map, a text, etc.

Our system will then select an application that uses the auto-radio of the car to transmit the feeds as audio news flashes to the user when driving. This is why our last presented application proposes an automatic RSS feeds trans-coder. News aggregators are widely used to read RSS feeds but they require the user to be in front of a screen. While moving, people usually do not have any display, or very small ones. Moreover, they need to perform actions to get access to the news: download a tool, choose to generate audio files from the news, and send them to e.g. an MP3 player. We propose a system that automatically detects when the user leaves the computer

room and directly sends the trans-coded news onto the user Smartphone. All the aggregated pieces of news are then transmitted to the user who can listen to them without any action.

In order to validate our engine, we made experimentations with these applications as test cases and provided a questionnaire to the users.

4.3 Validation questionnaire

Ten voluntary people, aged from 20 to 35 whose seven males, took part to the experimentation and answered to our questionnaire.

Once users tried our system, we proposed them a SUMI-like (i.e. Software Usability Measurement Inventory) questionnaire¹ [McSweeney 1992] to the users. Indeed, several types of validity studies [Ravden and Johnson 1989][Saunders and Jones 1992][Wong and Rengger 1990] have already been conducted with SUMI, whose one of them concerns laboratory-based studies (carried out in the Human Factors Research Group). This questionnaire is composed of two parts that we describe here.

The first part is composed of questions about the user profile: the age, gender, but also if the user is used to work with computer, within VR environment and VR equipment. This part concludes with questions about the training for using the system (availability and length) and if the time to use the system was also adequate.

The second part of the questionnaire is composed of fifty statements. The user must answer to all of them by marking one of the three proposed boxes labeled: “Disagree”, “Undecided”, and “Agree”. It is also firstly noticed that marking the “Undecided” boxes means that the user cannot make up his/her mind, or that the statement has no relevance to the software or the situation. Secondly, it is added that marking the “Disagree” or “agree” boxes does not necessarily indicate a strong disagreement (respectively agreement) but only a general feeling most of the time.

The questions of this second part concern various topics: responsiveness of the software, quality of proposed instructions, global satisfaction about the software, possible improvements, intuitiveness, and attractiveness of the software. In the next section, we will analyze the users’ answers and their behavior during the experiment.

5 Conclusion

In this paper, we wanted to show a rendering engine that automatically selects a rendering application depending on the user context, available devices and information type. We implemented such a system and evaluated it with the help of users, experimentations, and well-known questionnaires.

Our adapted SUMI questionnaire (results in Table 1), filled by every participant, first reveals us that our system is considered by every subject as “easy to configure” (creation of the user’s profile) even if most of them (8 people) also confessed they would maybe not know how to do it by themselves. However, length of setup was not considered as long and the possibility to save and load profiles was strongly appreciated.

Secondly, our questionnaires confirmed us that our software is fast to propose an adequate rendering application that perfectly fits with the requirements. But it also informs us about its limitations: As we did not implement all the test case applications, we were also limited with the choice of inputs. Nevertheless, when the application was made, it was well launched. In the other case, the system

	YES	Average	NO
Intuitiveness of GUI	0	8	2
Decision step’s length	10	0	0
Global system efficiency	7	3*	0

Table 1: Adaptation experiment average results (24 testers)

*Due to small number of test cases

simply showed on the terminal which type of devices could be used to render this information, e.g. “use a PDA with sound channel” if the user was blind and walking in the street.

The results seem to be very promising, and especially with the use of semantics on an ontology-based system. Indeed, our system can easily be extended by adding criteria in our ontology, and populated with new applications and devices.

Nevertheless, some improvements could be performed, particularly on the user interface at the setup (creation of the user’s profile). Actually, the setup interface is strictly based on the graphic channel (information to select on the screen with a mouse) and we did not plan that blind people could perform this step. And it would be also very interesting to populate the ontology with several rendering applications in order to do further validations.

Acknowledgements

This research has been partially supported by the European Coordination Action: FOCUS K3D (<http://www.focusk3d.eu>).

References

- AARTS, E. 2004. Ambient intelligence: A multimedia perspective. *IEEE MultiMedia* 11, 1, 12–19.
- COHEN, P. R., AND MCGEE, D. R. 2004. Tangible multimodal interfaces for safety-critical applications. *Commun. ACM* 47, 1, 41–46.
- COHEN, P. R., AND OVIATT, S. L. 1994. The role of voice in human-machine communication. 34–75.
- CORRADINI, A., MEHTA, M., BENSON, N., MARTIN, J., AND ABRILIAN, S. 2003. Multimodal input fusion in human computer interaction. In *NATO-ASI Conference on Data Fusion for Situation Monitoring, Incident Detection, Alert, and Response Management*.
- COUDERC, P., AND KERMARREC, A.-M. 1999. Improving level of service for mobile users using context-awareness. In *SRDS '99: Proceedings of the 18th IEEE Symposium on Reliable Distributed Systems*, IEEE Computer Society, Washington, DC, USA, 24.
- DIX, A., FINLAY, J., ABOWD, G., AND BEALE, R. 1997. *Human-computer interaction*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.
- FLIPPO, F., KREBS, A., AND MARSIC, I. 2003. A framework for rapid development of multimodal interfaces. In *ICMI '03: Proceedings of the 5th international conference on Multimodal interfaces*, ACM, New York, NY, USA, 109–116.
- GUI, F., ADJOUADI, M., AND RISHE, N. 2009. Personalized approach for mobile search. In *CSIE '09: Proceedings of the 2009 WRI World Congress on Computer Science and Information Engineering*, IEEE Computer Society, Washington, DC, USA, 322–326.

¹<http://sumi.ucc.ie/>

- GUNES, H., PICCARDI, M., AND JAN, T. 2004. Face and body gesture recognition for a vision-based multimodal analyzer. In *VIP '05: Proceedings of the Pan-Sydney area workshop on Visual information processing*, Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 19–28.
- HÖÖK, K. 1998. Tutorial 2: Designing and evaluating intelligent user interfaces. In *IUI '98: Proceedings of the 3rd international conference on Intelligent user interfaces*, ACM, New York, NY, USA, 5–6.
- JACQUEMIN, C., AND DE LAUBIER, S. 2006. Transmodal feedback as a new perspective for audio-visual effects. In *NIME '06: Proceedings of the 2006 conference on New interfaces for musical expression*, IRCAM — Centre Pompidou, Paris, France, France, 156–161.
- JAIMES, A., AND SEBE, N. 2007. Multimodal human-computer interaction: A survey. *Comput. Vis. Image Underst.* 108, 1-2, 116–134.
- KALTZ, J. W. 2006. Using context-awareness to support user interaction with web services. In *AICT-ICIW '06: Proceedings of the Advanced Int'l Conference on Telecommunications and Int'l Conference on Internet and Web Applications and Services*, IEEE Computer Society, Washington, DC, USA, 155.
- MALKAWI, A., AND SRINIVASAN, R. 2004. Multimodal human-computer interaction for immersive visualization: integrating speech-gesture recognition and augmented reality for indoor environments. In *International Association of Science and Technology for Development Conference on Computer Graphics and Imaging*.
1991. An agenda for human-computer interaction: science and engineering serving human needs. *SIGCHI Bull.* 23, 4, 17–32.
- MARTINEZ-RUIZ, F. J., VANDERDONCKT, J., AND ARTEAGA, JAIME MU N. 2008. Context-aware generation of user interface containers for mobile devices. In *ENC '08: Proceedings of the 2008 Mexican International Conference on Computer Science*, IEEE Computer Society, Washington, DC, USA, 63–72.
- MCSWEENEY, R. 1992. Sumi – a psychometric approach to software evaluation. Unpublished MA (Qual) thesis in Applied Psychology, University College Cork, Ireland.
- MOWAFI, Y., AND ZHANG, D. 2007. A user-centered approach to context-awareness in mobile computing. In *MOBIQUITOUS '07: Proceedings of the 2007 Fourth Annual International Conference on Mobile and Ubiquitous Systems: Networking & Services (MobiQuitous)*, IEEE Computer Society, Washington, DC, USA, 1–3.
- QVARFORDT, P., AND ZHAI, S. 2005. Conversing with the user based on eye-gaze patterns. In *CHI '05: Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, New York, NY, USA, 221–230.
- RAUTEK, P., AND BRUCKNER, S. 2007. Semantic layers for illustrative volume rendering. *IEEE Transactions on Visualization and Computer Graphics* 13, 6, 1336–1343. Member-Groller, Edward.
- RAVDEN, S., AND JOHNSON, G. 1989. *Evaluating usability of human-computer interfaces: a practical method*. Halsted Press, New York, NY, USA.
- SALAMIN, P., VEXO, F., AND THALMANN, D. In *ACM Symposium on Virtual Reality Software and Technology (VRST '06)*.
- SALAMIN, P., THALMANN, D., AND VEXO, F. 2007. Visualization Learning for Visually Impaired People. In *The 2nd International Conference of E-Learning and Games: Edutainment 2007*, 171–181.
- SALAMIN, P., THALMANN, D., AND VEXO, F. 2008. Improved Third-Person Perspective: a solution reducing occlusion of the 3PP? In *VRCAI 2008, the 7th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry*, 1–6.
- SALAMIN, P., THALMANN, D., VEXO, F., AND GIROUD, S. 2008. Two-arm haptic force-feedbacked aid for the shoulder and elbow telerehabilitation. In *Edutainment 2008*, 381–390.
- SALAMIN, P., THALMANN, D., AND VEXO, F. 2009. Intelligent switch: An algorithm to provide the best third-person perspective in augmented reality. In *22nd Annual Conference on Computer Animation and Social Agents (CASA 2009)*.
- SALAMIN, P., THALMANN, D., AND VEXO, F. 2009. Stay Tuned! an Automatic RSS feeds Trans-coder. In *3rd International Conference on Intelligent Technologies for Interactive Entertainment*.
- SAUNDERS, C. S., AND JONES, J. W. 1992. Measuring performance of the information systems function. *J. Manage. Inf. Syst.* 8, 4, 63–82.
- SHAMSFARD, M., AND ABDOLLAHZADEH BARFOROUSH, A. 2003. The state of the art in ontology learning: a framework for comparison. *Knowl. Eng. Rev.* 18, 4, 293–316.
- SIBERT, L. E., AND JACOB, R. J. K. 2000. Evaluation of eye gaze interaction. In *CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, New York, NY, USA, 281–288.
- SPYNS, P., TANG, Y., AND MEERSMAN, R. 2008. An ontology engineering methodology for dogma. *Appl. Ontol.* 3, 1-2, 13–39.
- SUTTERER, M., DROEGEHORN, O., AND DAVID, K. 2008. User profile selection by means of ontology reasoning. In *AICT '08: Proceedings of the 2008 Fourth Advanced International Conference on Telecommunications*, IEEE Computer Society, Washington, DC, USA, 299–304.
- THALMANN, D., SALAMIN, P., OTT, R., GUTIRREZ, M., AND VEXO, F. 2006. Advanced Mixed Reality Technologies for Surveillance and Risk Prevention Applications. In *21st International Symposium on Computer and Information Sciences (ISCIS '06)*, 13–23.
2005. Real-time gesture recognition by learning and selective control of visual interest points. *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 3, 351–364. Member-Kirishima, Toshiyuki and Member-Sato, Kosuke and Member-Chihara, Kunihiro.
- WONG, G., AND RENGGER, R. 1990. The validity of questionnaires designed to measure user-satisfaction of computer systems. Tech. rep., National Physical Laboratory report DITC 169/90, Teddington, Middx., UK.
- YE, J., COYLE, L., DOBSON, S., AND NIXON, P. 2007. Ontology-based models in pervasive computing systems. *Knowl. Eng. Rev.* 22, 4, 315–347.
- YOON, H., KIM, E., AND LEE, M. 2007. A user context awareness model for mobile environment processing. In *SERA '07: Proceedings of the 5th ACIS International Conference on Software Engineering Research, Management & Applications*, IEEE Computer Society, Washington, DC, USA, 293–296.